

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/516,697  
Source: PG 110  
Date Processed by STIC: 12/20/04

# *ENTERED*



PCT

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,697

TIME: 12:24:18

Input Set : A:\Avalon 163.txt

Output Set: N:\CRF4\12202004\J516697.raw

```

3 <110> APPLICANT: Avalon Pharmaceuticals
5 <120> TITLE OF INVENTION: Cancer-Linked Gene as Target for Chemotherapy
7 <130> FILE REFERENCE: 689290-163
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/516,697
C--> 10 <141> CURRENT FILING DATE: 2004-12-01
12 <150> PRIOR APPLICATION NUMBER: US/60/386,793
13 <151> PRIOR FILING DATE: 2002-06-07
15 <160> NUMBER OF SEQ ID NOS: 6
17 <170> SOFTWARE: PatentIn version 3.0
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 1669
21 <212> TYPE: DNA
22 <213> ORGANISM: Homo sapiens
24 <400> SEQUENCE: 1
25 tgtgagtcac caaggaaggc agcggcagct ccactcagcc agtaccacaga tacgctggga      60
26 accttcccca gccatggctt ccctggggca gatcctcttc tggagcataa ttagcatcat      120
27 cattattctg gctggagcaa ttgcactcat cattggcttt ggtatttcag aagtctctgt      180
28 ctggctttca gcaatgaagg gtttggtgtt agaagttcca aggcttccct tagcattgat      240
29 ctttgcttcc tgaactgcag ggagacactc catcacagtc actactgtcg cctcagctgg      300
30 gaacattggg gaggatggaa tcctgagctg cacttttgaa cctgacatca aactttctga      360
31 tatcgtgata caatggctga aggaagggtg tttaggcttg gtccatgagt tcaaagaagg      420
32 caaagatgag ctgtcggagc aggatgaaat gttcagaggc cggacagcag tgtttgctga      480
33 tcaagtgata gttggcaatg cctctttgct gctgaaaaac gtgcaactca cagatgctgg      540
34 cacctacaaa tggtatatca tcacttctaa aggcaagggg aatgctaacc ttgagtataa      600
35 aactggagcc ttcagcatgc cggaagtga tgtggactat aatgccagct cagagacctt      660
36 gcggtgtgag gctccccgat ggttcccca gccacagtg gtctgggcat cccaagttga      720
37 ccaggagacc aacttctcgg aagtctccaa taccagcttt gagctgaact ctgagaatgt      780
38 gaccatgaag gttgtgtctg tgctctacaa tgttacgac aacaacacat actcctgtat      840
39 gattgaaaat gacattgcca aagcaacagg ggatatcaaa gtgacagaat cggagatcaa      900
40 aaggcggagt cacctacagc tgctaaactc aaaggcttct ctgtgtgtct cttctttctt      960
41 tgccatcagc tgggcacttc tgccctctcag cccttacctg atgctaaaat aatgtgcctc     1020
42 ggccacaaaa aagcatgcaa agtcattggt acaacaggga tctacagaac tatttcacca     1080
43 ccagatatga cctagtttta tatttctggg aggaaatgaa ttcatatcta gaagtctgga     1140
44 gtgagcaaac aagagcaaga acaaaaaaga agccaaaagc agaaggctcc aatatgaaca     1200
45 agataaatct atcttcaaag acatattaga agttgggaaa ataattcatg tgaactagac     1260
46 aagtgtgtta agagtataa gtaaaatgca cgtggagaca agtgcacccc cagatctcag     1320
47 ggacctcccc ctgctgttca cctggggagt gagaggacag gatagtgcac gttctttgtc     1380
48 tctgaatttt tagttatatg tgctgtaatg ttgctctgag gaagccccctg gaaagtcctat     1440
49 cccaacatat ccacatctta tattccacaa attaatgtgt agtatgtacc ctaagacgct     1500
50 gctaattgac tgccacttcg caactcaggg gcggctgcat tttagtaatg ggtcaaataa     1560
51 ttcacttttt atgatgcttc caaagggtgcc ttggcttctc ttcccaactg acaaatagcca     1620
52 aagttgagaa aatgatcat aatttttagca taaacagagc agtcggcgga      1669
55 <210> SEQ ID NO: 2

```

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,697

TIME: 12:24:18

Input Set : A:\Avalon 163.txt

Output Set: N:\CRF4\12202004\J516697.raw

56 &lt;211&gt; LENGTH: 1004

57 &lt;212&gt; TYPE: DNA

58 &lt;213&gt; ORGANISM: Homo sapiens

60 &lt;400&gt; SEQUENCE: 2

```

61 cttcttttta aacaaacaaa tgcgggttta tttctcagat gatgttcac cgtgaatggt      60
62 ccagggaagg acctttcacc ttgtctatat ggcattatgt catcacaagc tctgaggctt      120
63 ctccctttcca tcctgcgtgg acagctaaga cctcagtttt caatagcatc tagagcagtg      180
64 ggactcagct ggggtgattt cgccccccat ctccggggga atgtctgaag acaattttgg      240
65 ttacctcaat gagggagtgg aggaggatac agtgctacta ccaactagtg gataaaggcc      300
66 agggatgctg ctcaacctcc taccatgtac aggacgtctc cccattacaa ctacccaatc      360
67 cgaagtgtca actgtgtcag gactaagaaa ccctggtttt gagtagaaaa gggcctggaa      420
68 agaggggagc caacaaatct gtctgcttcc tcacattagt cattggcaaa taagcattct      480
69 gtctcttttg ctgctgcctc agcacagaga gccagaactc tatcgggcac caggataaca      540
70 tctctcagtg aacagagttg acaaggccta tgggaaatgc ctgatggat tatcttcagc      600
71 ttgttgagct tctaagtttc tttcccttca ttctaccctg caagccaagt tctgtaagag      660
72 aaatgcctga gttctagctc aggtttttctt actctgaatt tagatctcca gacccttctt      720
73 ggccacaatt caaatataag caacaaacat ataccttcca tgaagcacac acagactttt      780
74 gaaagcaagg acaatgactg cttgaattga ggccttgagg aatgaagctt tgaaggaaaa      840
75 gaatactttg tttccagccc ccttcccaca ctcttcatgt gttaaccact gccttcctgg      900
76 accttgagc cagcgtgact gtattacatg ttgttataga aaactgattt tagagttctg      960
77 atcgttcaag agaatgatta aatatacatt tcctaaaaaa atgt                                     1004

```

80 &lt;210&gt; SEQ ID NO: 3

81 &lt;211&gt; LENGTH: 1808

82 &lt;212&gt; TYPE: DNA

83 &lt;213&gt; ORGANISM: Homo sapiens

85 &lt;400&gt; SEQUENCE: 3

```

86 tgtgagtcac caaggaaggc agcggcagct ccactcagcc agtaccaga tacgctggga      60
87 accttcccca gccatggctt ccctggggca gatcctcttc tggagcataa ttagcatcat      120
88 cattattctg gctggagcaa ttgcactcat cattggcttt ggtatttcag ggagacactc      180
89 catcacagtc actactgtcg ctcagctgg gaacattggg gaggatggaa tcctgagctg      240
90 cacttttgaa cctgacatca aactttctga tatcgtgata caatggctga aggaagggtg      300
91 tttaggcttg gtccatgagt tcaaagaagg caaagatgag ctgtcggagc aggatgaaat      360
92 gttcagaggc cggacagcag tgtttgctga tcaagtata gttggcaatg cctctttgctg      420
93 gctgaaaaac gtgcaactca cagatgctgg cacctacaaa tgttatatca tactttctaa      480
94 aggcaagggg aatgctaacc ttgagtataa aactggagcc ttcagcatgc cggaagtga      540
95 tgtggactat aatgccagct cagagacctt gcggtgtgag gctccccgat gggtcccca      600
96 gccacagtg gtctgggcat ccaagttga ccaggagcc aacttctcgg aagtctcaa      660
97 taccagcttt gagctgaact ctgagaatgt gaccatgaag gttgtgtctg tgctctacaa      720
98 tgttacgata aacaacacat actcctgtat gattgaaaat gacattgcca aagcaacagg      780
99 ggatatcaaa gtgacagaat cggagatcaa aaggcggagt cacctacagc tgctaaactc      840
100 aaaggcttct ctgtgtgtct cttctttctt tgccatcagc tgggcacttc tgctctcag      900
101 cccttacctg atgctaaaaa aatgtgcctc ggccacaaaa aagcatgcaa agtcattgtt      960
102 acaacaggga tctacagaac tatttcacca ccagatatga cctagtttta tatctctggg      1020
103 aggaaatgaa ttcatatcta gaagtctgga gtgagcaaac aagagcaaga aacaaaaaga      1080
104 agccaaaagc agaaggctcc aatatgaaca agataaatct atcttcaaag acatattaga      1140
105 agttgggaaa ataattcatg tgaactagat gtcaactgtg tcaggactaa gaaaccctgg      1200
106 ttttagtagt aaaagggcct ggaaagaggg gagccaaca atctgtctgc ttctcacat      1260
107 tagtcaattg caaataagca ttctgtctct ttggctgctg cctcagcaca gagagccaga      1320
108 actctatcgg gcaccaggat aacatctctc agtgaaacaga gttgacaagg cctatgggaa      1380

```

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,697

TIME: 12:24:18

Input Set : A:\Avalon 163.txt

Output Set: N:\CRF4\12202004\J516697.raw

```

109 atgcctgatg ggattatctt cagcttggtg agcttctaag tttctttccc ttcattctac 1440
110 cctgcaagcc aagttctgta agagaaatgc ctgagttcta gctcaggttt tcttactctg 1500
111 aatttagatc tccagaccct tcctggccac aattcaaatt aaggcaacaa acatatacct 1560
112 tccatgaagc acacacagac ttttgaaagc aaggacaatg actgcttgaa ttgaggcctt 1620
113 gaggaatgaa gctttgaagg aaaagaatac tttgtttcca gcccccttcc cacactcttc 1680
114 atgtgttaac cactgccttc ctggaccttg gagccacggt gactgtatta catgttggtta 1740
115 tagaaaactg atttttagagt tctgatcggt caagagaatg attaaatata catttcctaa 1800
116 aaaaatgt 1808
119 <210> SEQ ID NO: 4
120 <211> LENGTH: 1898
121 <212> TYPE: DNA
122 <213> ORGANISM: Homo sapiens
124 <400> SEQUENCE: 4
125 tgtgagtcac caaggaaggc agcggcagct ccactcagcc agtaccacaga tacgctggga 60
126 accttccccca gccatggctt ccctggggca gatcctcttc tggagcataa ttagcatcat 120
127 cattattctg gctggagcaa ttgcactcat cattggcttt ggtatttcag aagtctctgt 180
128 ctggctttca gcaatgaagg gtttggttgt agaagttcca aggcttccct tagcattgat 240
129 ctttgcttcc tgaactgcag ggagacactc catcacagtc actactgtcg cctcagctgg 300
130 gaacattggg gaggatggaa tcctgagctg cacttttgaa cctgacatca aactttctga 360
131 tatcgtgata caatggctga aggaagggtg tttaggcttg gtccatgagt tcaaagaagg 420
132 caaagatgag ctgtcggagc aggatgaaat gtccagaggc cggacagcag tgtttgctga 480
133 tcaagtgata gttggcaatg cctctttgct gctgaaaaac gtgcaactca cagatgctgg 540
134 cacctacaaa tgttatatca tcacttctaa aggcaagggg aatgctaacc ttgagtataa 600
135 aactggagcc ttcagcatgc cggaaagtga tgtggactat aatgccagct cagagacctt 660
136 gcggtgtgag gctccccgat ggttccccca gccacagtg gtctgggcat cccaagttga 720
137 ccaggagcc aacttctcgg aagtctccaa taccagcttt gagctgaact ctgagaatgt 780
138 gaccatgaag gttgtgtctg tgctctacaa tgttacgata aacaacacat actcctgtat 840
139 gattgaaaat gacattgcca aagcaacagg ggatacaaaa gtgacagaat cggagatcaa 900
140 aaggcggagt cacctacagc tgctaaactc aaaggcttct ctgtgtgtct cttctttctt 960
141 tgccatcagc tgggcacttc tgctctcag cccttacctg atgctaaaaat aatgtgcctc 1020
142 ggccacaaaa aagcatgcaa agtcattgtt acaacaggga tctacagaaac tatttcacca 1080
143 ccagatatga cctagtttta ttttctggg aggaaatgaa ttcatatcta gaagtctgga 1140
144 gtgagcaaac aagagcaaga acaaaaaaga agccaaaagc agaaggctcc aatatgaaca 1200
145 agataaatct atcttcaaag acatattaga agttgggaaa ataattcatg tgaactagat 1260
146 gtcaactgtg tcaggactaa gaaaccctgg ttttgagtag aaaagggcct ggaaagaggg 1320
147 gagccaacaa atctgtctgc ttcctcacat tagtcattgg caaataagca ttctgtctct 1380
148 ttggctgctg cctcagcaca gagagccaga actctatcgg gcaccaggat aacatctctc 1440
149 agtgaacaga gttgacaagg cctatgggaa atgcctgatg ggattatctt cagcttggtg 1500
150 agcttctaag tttctttccc ttcattctac cctgcaagcc aagttctgta agagaaatgc 1560
151 ctgagttcta gctcaggttt tcttactctg aatttagatc tccagaccct tcctggccac 1620
152 aattcaaatt aaggcaacaa acatatacct tccatgaagc acacacagac ttttgaaagc 1680
153 aaggacaatg actgcttgaa ttgaggcctt gaggaatgaa gctttgaagg aaaagaatac 1740
154 tttgtttcca gcccccttcc cacactcttc atgtgttaac cactgccttc ctggaccttg 1800
155 gagccacggt gactgtatta catgttggtta tagaaaactg atttttagagt tctgatcggt 1860
156 caagagaatg attaaatata catttcctaa aaaaatgt 1898
159 <210> SEQ ID NO: 5
160 <211> LENGTH: 336
161 <212> TYPE: PRT
162 <213> ORGANISM: Homo sapiens

```

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,697

TIME: 12:24:18

Input Set : A:\Avalon 163.txt

Output Set: N:\CRF4\12202004\J516697.raw

```

164 <400> SEQUENCE: 5
165 Val Ser His Gln Gly Arg Gln Arg Gln Leu His Ser Ala Ser Thr Gln
166 1 5 10 15
168 Ile Arg Trp Glu Pro Ser Pro Ala Met Ala Ser Leu Gly Gln Ile Leu
169 20 25 30
171 Phe Trp Ser Ile Ile Ser Ile Ile Ile Leu Ala Gly Ala Ile Ala
172 35 40 45
174 Leu Ile Ile Gly Phe Gly Ile Ser Glu Val Ser Val Trp Leu Ser Ala
175 50 55 60
177 Met Lys Gly Leu Val Val Glu Val Pro Arg Leu Pro Leu Ala Leu Ile
178 65 70 75 80
180 Phe Ala Ser Cys Thr Ala Gly Arg His Ser Ile Thr Val Thr Thr Val
181 85 90 95
183 Ala Ser Ala Gly Asn Ile Gly Glu Asp Gly Ile Leu Ser Cys Thr Phe
184 100 105 110
186 Glu Pro Asp Ile Lys Leu Ser Asp Ile Val Ile Gln Trp Leu Lys Glu
187 115 120 125
189 Gly Val Leu Gly Leu Val His Glu Phe Lys Glu Gly Lys Asp Glu Leu
190 130 135 140
192 Ser Glu Gln Asp Glu Met Phe Arg Gly Arg Thr Ala Val Phe Ala Asp
193 145 150 155 160
195 Gln Val Ile Val Gly Asn Ala Ser Leu Arg Leu Lys Asn Val Gln Leu
196 165 170 175
198 Thr Asp Ala Gly Thr Tyr Lys Cys Tyr Ile Ile Thr Ser Lys Gly Lys
199 180 185 190
201 Gly Asn Ala Asn Leu Glu Tyr Lys Thr Gly Ala Phe Ser Met Pro Glu
202 195 200 205
204 Val Asn Val Asp Tyr Asn Ala Ser Ser Glu Thr Leu Arg Cys Glu Ala
205 210 215 220
207 Pro Arg Trp Phe Pro Gln Pro Thr Val Val Trp Ala Ser Gln Val Asp
208 225 230 235 240
210 Gln Gly Ala Asn Phe Ser Glu Val Ser Asn Thr Ser Phe Glu Leu Asn
211 245 250 255
213 Ser Glu Asn Val Thr Met Lys Val Val Ser Val Leu Tyr Asn Val Thr
214 260 265 270
216 Ile Asn Asn Thr Tyr Ser Cys Met Ile Glu Asn Asp Ile Ala Lys Ala
217 275 280 285
219 Thr Gly Asp Ile Lys Val Thr Glu Ser Glu Ile Lys Arg Arg Ser His
220 290 295 300
222 Leu Gln Leu Leu Asn Ser Lys Ala Ser Leu Cys Val Ser Ser Phe Phe
223 305 310 315 320
225 Ala Ile Ser Trp Ala Leu Leu Pro Leu Ser Pro Tyr Leu Met Leu Lys
226 325 330 335
229 <210> SEQ ID NO: 6
230 <211> LENGTH: 306
231 <212> TYPE: PRT
232 <213> ORGANISM: Homo sapiens
234 <400> SEQUENCE: 6
235 Val Ser His Gln Gly Arg Gln Arg Gln Leu His Ser Ala Ser Thr Gln

```

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,697

TIME: 12:24:18

Input Set : A:\Avalon 163.txt

Output Set: N:\CRF4\12202004\J516697.raw

```

236 1           5           10           15
238 Ile Arg Trp Glu Pro Ser Pro Ala Met Ala Ser Leu Gly Gln Ile Leu
239           20           25           30
241 Phe Trp Ser Ile Ile Ser Ile Ile Ile Ile Leu Ala Gly Ala Ile Ala
242           35           40           45
244 Leu Ile Ile Gly Phe Gly Ile Ser Gly Arg His Ser Ile Thr Val Thr
245           50           55           60
247 Thr Val Ala Ser Ala Gly Asn Ile Gly Glu Asp Gly Ile Leu Ser Cys
248 65           70           75           80
250 Thr Phe Glu Pro Asp Ile Lys Leu Ser Asp Ile Val Ile Gln Trp Leu
251           85           90           95
253 Lys Glu Gly Val Leu Gly Leu Val His Glu Phe Lys Glu Gly Lys Asp
254           100          105          110
256 Glu Leu Ser Glu Gln Asp Glu Met Phe Arg Gly Arg Thr Ala Val Phe
257           115          120          125
259 Ala Asp Gln Val Ile Val Gly Asn Ala Ser Leu Arg Leu Lys Asn Val
260           130          135          140
262 Gln Leu Thr Asp Ala Gly Thr Tyr Lys Cys Tyr Ile Ile Thr Ser Lys
263 145           150          155          160
265 Gly Lys Gly Asn Ala Asn Leu Glu Tyr Lys Thr Gly Ala Phe Ser Met
266           165          170          175
268 Pro Glu Val Asn Val Asp Tyr Asn Ala Ser Ser Glu Thr Leu Arg Cys
269           180          185          190
271 Glu Ala Pro Arg Trp Phe Pro Gln Pro Thr Val Val Trp Ala Ser Gln
272           195          200          205
274 Val Asp Gln Gly Ala Asn Phe Ser Glu Val Ser Asn Thr Ser Phe Glu
275           210          215          220
277 Leu Asn Ser Glu Asn Val Thr Met Lys Val Val Ser Val Leu Tyr Asn
278 225           230          235          240
280 Val Thr Ile Asn Asn Thr Tyr Ser Cys Met Ile Glu Asn Asp Ile Ala
281           245          250          255
283 Lys Ala Thr Gly Asp Ile Lys Val Thr Glu Ser Glu Ile Lys Arg Arg
284           260          265          270
286 Ser His Leu Gln Leu Leu Asn Ser Lys Ala Ser Leu Cys Val Ser Ser
287           275          280          285
289 Phe Phe Ala Ile Ser Trp Ala Leu Leu Pro Leu Ser Pro Tyr Leu Met
290           290          295          300
292 Leu Lys
293 305

```

VERIFICATION SUMMARY

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,697

TIME: 12:24:19

Input Set : A:\Avalon 163.txt

Output Set: N:\CRF4\12202004\J516697.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date